

statistical system. Joining me in this discussion were Joe Daly and Ed Goldfield of the Bureau of the Census, Bob Steffes and Rudy Mendelssohn of BLS and Ezra Glaser of the Patent Office. In a second initiative Mendelssohn undertook a more intensive study based upon the earlier survey of machine readable records conducted by OSS to try to get a better fix on what it will take to establish an archival function. He in turn was assisted by many people in the agencies. Lastly, the National Bureau of Standards was used as a vehicle to assemble a small group of knowledgeable people in an attempt to specify more clearly the essential elements of a data service center required to provide a range of facilitating services. Ezra Glaser, Marshall Wood, and Dave Rosenblatt were the principal contributors to this effort although conversations included Sam Alexander and other members of his staff. Paul Krueger and I also participated in this effort.

In addition, I have engaged in many discussions of substantive issues with a number of knowledgeable people in the Federal agencies (both statistical agencies and program agencies) and in the universities.

I particularly want to acknowledge the invaluable assistance that I have received from Paul Krueger on your staff. He has given me continuous support, assistance and encouragement. He has joined me in many of the meetings and discussions with the aforementioned and has made his own valuable contribution to the thinking process.

The form and content of the report, of course, remains my own responsibility. I believe that the general conclusions and recommendations are sound and supported in whole or in part by the informed judgment of many others beside myself, but I do not attribute the views of this report specifically to any one or all of its many contributors.

In writing the report I have incorporated material included in earlier memorandums as well as sections that contain explanation and argument that is superfluous from the point of view of the informed staff member of the Office of Statistical Standards. I did so because I assumed that this report might be used in whole or in part to communicate elements of this problem and the recommended solutions to more than one group. I attempted, therefore, to include a comprehensive discussion of the problems and opportunities.

The report makes clear that my own understanding and evaluation of this problem has modified somewhat in the course of the study. I now feel that the production standards and practices are a more important element in both problems and solution. This, as well as other considerations, leads me to be less sanguine about the possibility or the desirability of keeping the issues of organization in the background. I think that there might be some benefit in our discussing this and several related issues on an informal basis.

Let me say that I have enjoyed working with the Office of Statistical Standards on this problem. I hope that the results are constructive in serving your needs and objectives.

Sincerely yours,

EDGAR S. DUNN, Jr.

SUMMARY

The Ruggles Committee report recommending the establishment of a National Data Center is only one of the more manifest expressions of concern, dissatisfaction and frustration that have been surfacing among the groups that use numerical records for research, planning or decisionmaking at all levels. The problems at issue go far beyond the forms of discontent generated by special interests or marginal interests not served by public policy. They result from major changes on both the demand and supply side of the information process since World War II. Many people in this wider circle are attached to a rather naive data bank concept of the solution that does not incorporate an adequate appreciation of the basic problems in data use and data generation.

THE PROBLEM

The central problem of data use is one of associating numerical records and the greatest deficiency of the existing Federal statistical system is its failure to provide access to data in a way that permits the association of the elements of data sets in order to identify and measure the interrelationship among interdependent or related observations. This is true at virtually all levels of use and for all purposes from academic model builders to business market researchers.